

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 09/872,505A  
Source: 1FW16  
Date Processed by STIC: 10/3/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:50

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw

4 <110> APPLICANT: Andersen, Peter  
 5 Skjot, Rikke Louise Vinther  
 6 Okkels, Li Mei Meng  
 7 Brock, Inger  
 8 Oettinger, Thomas  
 10 <120> TITLE OF INVENTION: Nucleic Acid Fragments and Polypeptide Fragments Derived  
 from M.  
 11 Tuberculosis  
 13 <130> FILE REFERENCE: 670001-2002.6  
 15 <140> CURRENT APPLICATION NUMBER: 09/872,505A  
 16 <141> CURRENT FILING DATE: 2001-06-01  
 18 <150> PRIOR APPLICATION NUMBER: 09/804,980  
 19 <151> PRIOR FILING DATE: 2001-03-13  
 21 <150> PRIOR APPLICATION NUMBER: US09/615,947  
 22 <151> PRIOR FILING DATE: 2000-07-13  
 24 <150> PRIOR APPLICATION NUMBER: US09/246,191  
 25 <151> PRIOR FILING DATE: 1998-12-30  
 27 <150> PRIOR APPLICATION NUMBER: 60/070,488  
 28 <151> PRIOR FILING DATE: 1998-01-05  
 30 <150> PRIOR APPLICATION NUMBER: 60/144,011  
 31 <151> PRIOR FILING DATE: 1999-07-15  
 33 <150> PRIOR APPLICATION NUMBER: PCT/DK00/00398  
 34 <151> PRIOR FILING DATE: 2000-07-13  
 36 <160> NUMBER OF SEQ ID NOS: 68  
 38 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
 40 <210> SEQ ID NO: 1  
 41 <211> LENGTH: 324  
 42 <212> TYPE: DNA  
 43 <213> ORGANISM: Mycobacterium tuberculosis  
 45 <220> FEATURE:  
 46 <221> NAME/KEY: CDS  
 47 <222> LOCATION: (1)...(321)  
 49 <400> SEQUENCE: 1  
 50 ttg acc cac aag cgc act aaa cgc cag cca gcc atc gcc gca ggg ctc 48  
 51 Leu Thr His Lys Arg Thr Lys Arg Gln Pro Ala Ile Ala Ala Gly Leu  
 52 1 5 10 15  
 54 aac gcc ccg cgt cgg aat cgc gtt ggg cgg caa cat ggt tgg ccg gcc 96  
 55 Asn Ala Pro Arg Arg Asn Arg Val Gly Arg Gln His Gly Trp Pro Ala  
 56 20 25 30  
 58 gac gtt ccg tcc gcc gag cag cgc gcc caa cgg cag cgc gac ctc 144  
 59 Asp Val Pro Ser Ala Glu Gln Arg Arg Ala Gln Arg Gln Arg Asp Leu  
 60 35 40 45  
 62 gag gct atc cgc cga gcg tac gcc gag atg gtg gcg aca tca cac gaa 192  
 63 Glu Ala Ile Arg Arg Ala Tyr Ala Glu Met Val Ala Thr Ser His Glu

## RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:50

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw

```

64      50      55      60
66 atc gac gac gac aca gcc gaa ctg gcg ctg ttg tcg atg cat ctc gac      240
67 Ile Asp Asp Asp Thr Ala Glu Leu Ala Leu Leu Ser Met His Leu Asp
68 65      70      75      80
70 gat gag cag cgc cgg ctt gag gcg ggg atg aag ctc ggc tgg cat ccg      288
71 Asp Glu Gln Arg Arg Leu Glu Ala Gly Met Lys Leu Gly Trp His Pro
72      85      90      95
74 tat cac ttc ccc gac gaa ccc gac agc aaa cag tga      324
75 Tyr His Phe Pro Asp Glu Pro Asp Ser Lys Gln
76      100      105
78 <210> SEQ ID NO: 2
79 <211> LENGTH: 107
80 <212> TYPE: PRT
81 <213> ORGANISM: Mycobacterium tuberculosis
83 <400> SEQUENCE: 2
84 Met Thr His Lys Arg Thr Lys Arg Gln Pro Ala Ile Ala Ala Gly Leu
85 1      5      10      15
86 Asn Ala Pro Arg Arg Asn Arg Val Gly Arg Gln His Gly Trp Pro Ala
87      20      25      30
88 Asp Val Pro Ser Ala Glu Gln Arg Arg Ala Gln Arg Gln Arg Asp Leu
89      35      40      45
90 Glu Ala Ile Arg Arg Ala Tyr Ala Glu Met Val Ala Thr Ser His Glu
91      50      55      60
92 Ile Asp Asp Asp Thr Ala Glu Leu Ala Leu Leu Ser Met His Leu Asp
93 65      70      75      80
94 Asp Glu Gln Arg Arg Leu Glu Ala Gly Met Lys Leu Gly Trp His Pro
95      85      90      95
96 Tyr His Phe Pro Asp Glu Pro Asp Ser Lys Gln
97      100      105
99 <210> SEQ ID NO: 3
100 <211> LENGTH: 246
101 <212> TYPE: DNA
102 <213> ORGANISM: Mycobacterium tuberculosis
104 <220> FEATURE:
105 <221> NAME/KEY: CDS
106 <222> LOCATION: (1)...(243)
108 <400> SEQUENCE: 3
109 atg agc ggc cac gcg ttg gct gct cgg acg ttg ctg gcc gcc gcg gac      48
110 Met Ser Gly His Ala Leu Ala Ala Arg Thr Leu Leu Ala Ala Ala Asp
111 1      5      10      15
113 gag ctt gtc ggc ggc ccg cca gtc gag gct tcg gcc gcc gcg ctg gcc      96
114 Glu Leu Val Gly Gly Pro Pro Val Glu Ala Ser Ala Ala Ala Leu Ala
115      20      25      30
117 ggc gac gcc gcg ggc gca tgg cgg acc gcg gcc gtc gag ctt gcg cga      144
118 Gly Asp Ala Ala Gly Ala Trp Arg Thr Ala Ala Val Glu Leu Ala Arg
119      35      40      45
121 gcg ttg gtc cgc gct gtg gcg gag tcg cac ggc gtc gcg gcc gtt ttg      192
122 Ala Leu Val Arg Ala Val Ala Glu Ser His Gly Val Ala Ala Val Leu
123      50      55      60

```

## RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:50

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw

```

125 ttc gcc gcg acg gcc gcc gcg gcg gcg gcc gtc gac cgg ggt gat ccg      240
126 Phe Ala Ala Thr Ala Ala Ala Ala Val Asp Arg Gly Asp Pro
127 65                               70                               75                               80
129 ccg tga                                                                246
130 Pro
133 <210> SEQ ID NO: 4
134 <211> LENGTH: 81
135 <212> TYPE: PRT
136 <213> ORGANISM: Mycobacterium tuberculosis
138 <400> SEQUENCE: 4
139 Met Ser Gly His Ala Leu Ala Ala Arg Thr Leu Leu Ala Ala Ala Asp
140 1                               5                               10                               15
141 Glu Leu Val Gly Gly Pro Pro Val Glu Ala Ser Ala Ala Ala Leu Ala
142                               20                               25                               30
143 Gly Asp Ala Ala Gly Ala Trp Arg Thr Ala Ala Val Glu Leu Ala Arg
144                               35                               40                               45
145 Ala Leu Val Arg Ala Val Ala Glu Ser His Gly Val Ala Ala Val Leu
146                               50                               55                               60
147 Phe Ala Ala Thr Ala Ala Ala Ala Ala Val Asp Arg Gly Asp Pro
148 65                               70                               75                               80
149 Pro
151 <210> SEQ ID NO: 5
152 <211> LENGTH: 1200
153 <212> TYPE: DNA
154 <213> ORGANISM: Mycobacterium tuberculosis
156 <220> FEATURE:
157 <221> NAME/KEY: CDS
158 <222> LOCATION: (28)...(1140)
160 <400> SEQUENCE: 5
161 taataggcccc ccaacacatc ggagggga gtg atc acc atg ctg tgg cac gca atg      54
162                               Met Ile Thr Met Leu Trp His Ala Met
163                               1                               5
165 cca ccg gag cta aat acc gca cgg ctg atg gcc ggc gcg ggt ccg gct      102
166 Pro Pro Glu Leu Asn Thr Ala Arg Leu Met Ala Gly Ala Gly Pro Ala
167 10                               15                               20                               25
169 cca atg ctt gcg gcg gcc gcg gga tgg cag acg ctt tcg gcg gct ctg      150
170 Pro Met Leu Ala Ala Ala Ala Gly Trp Gln Thr Leu Ser Ala Ala Leu
171                               30                               35                               40
173 gac gct cag gcc gtc gag ttg acc gcg cgc ctg aac tct ctg gga gaa      198
174 Asp Ala Gln Ala Val Glu Leu Thr Ala Arg Leu Asn Ser Leu Gly Glu
175                               45                               50                               55
177 gcc tgg act gga ggt ggc agc gac aag gcg ctt gcg gct gca acg ccg      246
178 Ala Trp Thr Gly Gly Gly Ser Asp Lys Ala Leu Ala Ala Ala Thr Pro
179                               60                               65                               70
181 atg gtg gtc tgg cta caa acc gcg tca aca cag gcc aag acc cgt gcg      294
182 Met Val Val Trp Leu Gln Thr Ala Ser Thr Gln Ala Lys Thr Arg Ala
183                               75                               80                               85
185 atg cag gcg acg gcg caa gcc gcg gca tac acc cag gcc atg gcc acg      342
186 Met Gln Ala Thr Ala Gln Ala Ala Ala Tyr Thr Gln Ala Met Ala Thr

```

## RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:50

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw

187	90	95	100	105	
189	acg ccg tcg ctg ccg gag atc gcc gcc aac cac atc acc cag gcc gtc	390			
190	Thr Pro Ser Leu Pro Glu Ile Ala Ala Asn His Ile Thr Gln Ala Val				
191	110	115	120		
193	ctt acg gcc acc aac ttc ttc ggt atc aac acg atc ccg atc gcg ttg	438			
194	Leu Thr Ala Thr Asn Phe Phe Gly Ile Asn Thr Ile Pro Ile Ala Leu				
195	125	130	135		
197	acc gag atg gat tat ttc atc cgt atg tgg aac cag gca gcc ctg gca	486			
198	Thr Glu Met Asp Tyr Phe Ile Arg Met Trp Asn Gln Ala Ala Leu Ala				
199	140	145	150		
201	atg gag gtc tac cag gcc gag acc gcg gtt aac acg ctt ttc gag aag	534			
202	Met Glu Val Tyr Gln Ala Glu Thr Ala Val Asn Thr Leu Phe Glu Lys				
203	155	160	165		
205	ctc gag ccg atg gcg tcg atc ctt gat ccc ggc gcg agc cag agc acg	582			
206	Leu Glu Pro Met Ala Ser Ile Leu Asp Pro Gly Ala Ser Gln Ser Thr				
207	170	175	180	185	
209	acg aac ccg atc ttc gga atg ccc tcc cct ggc agc tca aca ccg gtt	630			
210	Thr Asn Pro Ile Phe Gly Met Pro Ser Pro Gly Ser Ser Thr Pro Val				
211	190	195	200		
213	ggc cag ttg ccg ccg gcg gct acc cag acc ctc ggc caa ctg ggt gag	678			
214	Gly Gln Leu Pro Pro Ala Ala Thr Gln Thr Leu Gly Gln Leu Gly Glu				
215	205	210	215		
217	atg agc ggc ccg atg cag cag ctg acc cag ccg ctg cag cag gtg acg	726			
218	Met Ser Gly Pro Met Gln Gln Leu Thr Gln Pro Leu Gln Gln Val Thr				
219	220	225	230		
221	tcg ttg ttc agc cag gtg ggc ggc acc ggc ggc ggc aac cca gcc gac	774			
222	Ser Leu Phe Ser Gln Val Gly Gly Thr Gly Gly Gly Asn Pro Ala Asp				
223	235	240	245		
225	gag gaa gcc gcg cag atg ggc ctg ctc ggc acc agt ccg ctg tcg aac	822			
226	Glu Glu Ala Ala Gln Met Gly Leu Leu Gly Thr Ser Pro Leu Ser Asn				
227	250	255	260	265	
229	cat ccg ctg gct ggt gga tca ggc ccc agc gcg ggc gcg ggc ctg ctg	870			
230	His Pro Leu Ala Gly Gly Ser Gly Pro Ser Ala Gly Ala Gly Leu Leu				
231	270	275	280		
233	cgc gcg gag tcg cta cct ggc gca ggt ggg tcg ttg acc cgc acg ccg	918			
234	Arg Ala Glu Ser Leu Pro Gly Ala Gly Ser Leu Thr Arg Thr Pro				
235	285	290	295		
237	ctg atg tct cag ctg atc gaa aag ccg gtt gcc ccc tcg gtg atg ccg	966			
238	Leu Met Ser Gln Leu Ile Glu Lys Pro Val Ala Pro Ser Val Met Pro				
239	300	305	310		
241	gcg gct gct gcc gga tcg tcg gcg acg ggt ggc gcc gct ccg gtg ggt	1014			
242	Ala Ala Ala Ala Gly Ser Ser Ala Thr Gly Gly Ala Ala Pro Val Gly				
243	315	320	325		
245	gcg gga gcg atg ggc cag ggt gcg caa tcc ggc ggc tcc acc agg ccg	1062			
246	Ala Gly Ala Met Gly Gln Gly Ala Gln Ser Gly Gly Ser Thr Arg Pro				
247	330	335	340	345	
249	ggt ctg gtc gcg ccg gca ccg ctc gcg cag gag cgt gaa gaa gac gac	1110			
250	Gly Leu Val Ala Pro Ala Pro Leu Ala Gln Glu Arg Glu Glu Asp Asp				
251	350	355	360		

## RAW SEQUENCE LISTING

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:50

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw

```

253 gag gac gac tgg gac gaa gag gac gac tgg tgagctcccg taatgacaac      1160
254 Glu Asp Asp Trp Asp Glu Glu Asp Asp Trp
255      365                      370
257 agacttcccg gccacccggg ccggaagact tgccaacatt      1200
259 <210> SEQ ID NO: 6
260 <211> LENGTH: 371
261 <212> TYPE: PRT
262 <213> ORGANISM: Mycobacterium tuberculosis
264 <400> SEQUENCE: 6
265 Met Ile Thr Met Leu Trp His Ala Met Pro Pro Glu Leu Asn Thr Ala
266 1      5      10      15
267 Arg Leu Met Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Ala
268      20      25      30
269 Gly Trp Gln Thr Leu Ser Ala Ala Leu Asp Ala Gln Ala Val Glu Leu
270      35      40      45
271 Thr Ala Arg Leu Asn Ser Leu Gly Glu Ala Trp Thr Gly Gly Gly Ser
272      50      55      60
273 Asp Lys Ala Leu Ala Ala Thr Pro Met Val Val Trp Leu Gln Thr
274 65      70      75      80
275 Ala Ser Thr Gln Ala Lys Thr Arg Ala Met Gln Ala Thr Ala Gln Ala
276      85      90      95
277 Ala Ala Tyr Thr Gln Ala Met Ala Thr Thr Pro Ser Leu Pro Glu Ile
278      100     105     110
279 Ala Ala Asn His Ile Thr Gln Ala Val Leu Thr Ala Thr Asn Phe Phe
280      115     120     125
281 Gly Ile Asn Thr Ile Pro Ile Ala Leu Thr Glu Met Asp Tyr Phe Ile
282      130     135     140
283 Arg Met Trp Asn Gln Ala Ala Leu Ala Met Glu Val Tyr Gln Ala Glu
284 145     150     155     160
285 Thr Ala Val Asn Thr Leu Phe Glu Lys Leu Glu Pro Met Ala Ser Ile
286      165     170     175
287 Leu Asp Pro Gly Ala Ser Gln Ser Thr Thr Asn Pro Ile Phe Gly Met
288      180     185     190
289 Pro Ser Pro Gly Ser Ser Thr Pro Val Gly Gln Leu Pro Pro Ala Ala
290      195     200     205
291 Thr Gln Thr Leu Gly Gln Leu Gly Glu Met Ser Gly Pro Met Gln Gln
292      210     215     220
293 Leu Thr Gln Pro Leu Gln Gln Val Thr Ser Leu Phe Ser Gln Val Gly
294 225     230     235     240
295 Gly Thr Gly Gly Gly Asn Pro Ala Asp Glu Glu Ala Ala Gln Met Gly
296      245     250     255
297 Leu Leu Gly Thr Ser Pro Leu Ser Asn His Pro Leu Ala Gly Gly Ser
298      260     265     270
299 Gly Pro Ser Ala Gly Ala Gly Leu Leu Arg Ala Glu Ser Leu Pro Gly
300      275     280     285
301 Ala Gly Gly Ser Leu Thr Arg Thr Pro Leu Met Ser Gln Leu Ile Glu
302      290     295     300
303 Lys Pro Val Ala Pro Ser Val Met Pro Ala Ala Ala Gly Ser Ser
304 305     310     315     320

```

**VERIFICATION SUMMARY**

DATE: 10/03/2005

PATENT APPLICATION: US/09/872,505A

TIME: 16:57:51

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10032005\I872505A.raw